

# Towards an immersive user interface for waypoint navigation of mobile robots

Greg Baker

*Bristol Robotics Laboratory*

*University of the West of England*

Bristol, UK

gregory.baker@brl.ac.uk

Paul Bremner

*Bristol Robotics Laboratory*

*University of the West of England*

Bristol, UK

paul.bremner@brl.ac.uk

Tom Bridgwater

*Bristol Robotics Laboratory*

*University of the West of England*

Bristol, UK

tom.bridgwater@uwe.ac.uk

Manuel Giuliani

*Bristol Robotics Laboratory*

*University of the West of England*

Bristol, UK

manuel.giuliani@brl.ac.uk

## ABSTRACT

In this paper, we investigate the utility of head-mounted display (HMD) interfaces for navigation of mobile robots. We focus on the selection of waypoint positions for the robot, whilst maintaining an egocentric view of the robot's environment. Inspired by virtual reality (VR) gaming, we propose a target selection method that uses the 6 degrees-of-freedom tracked controllers of a commercial VR headset. This allows an operator to point to the desired target position, in the vicinity of the robot, which the robot then autonomously navigates towards. A user study (37 participants) was conducted to examine the efficacy of this control strategy when compared to direct control, both with and without a communication delay. The results of the experiment showed that participants were able to learn how to use the novel system quickly, and the majority of participants reported a preference for waypoint control. Across all recorded metrics (task performance, operator workload and usability) the proposed waypoint control interface was not significantly affected by the communication delay, in contrast to direct control. The simulated experiment indicated that a real-world implementation of the proposed interface could be effective, but also highlighted the need to manage the negative effects of HMDs - particularly VR sickness.

## KEYWORDS

Mobile robots, teleoperation, immersive interfaces, supervisory control, waypoint control, user interface design, latency compensation, user study

### ACM Reference Format:

Greg Baker, Tom Bridgwater, Paul Bremner, and Manuel Giuliani. 2019. Towards an immersive user interface for waypoint navigation of mobile robots. In *Proceedings of VAM-HRI 2020*. ACM, New York, NY, USA, 9 pages.

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*VAM-HRI 2020, 23<sup>rd</sup> Mar, 2020, Cambridge, UK*

© 2019 Copyright held by the owner/author(s).

## 1 INTRODUCTION

Autonomy is the capability of a system to carry out tasks without human control [4]. Research has shown that systems with a higher level of autonomy can reduce operator workload and stress [34] and increase task performance [30], especially in the presence of communication delays between the operator and the the robot [7, 19]. However, the complexity of the tasks and/or environment can limit the level of autonomy that can be achieved. Furthermore, a lower level of autonomy may be deemed preferable if human involvement in the exploration/navigation process is one of the primary functions of the robot (e.g. telepresence robots). Often, the optimum level of autonomy lies somewhere between the two extremes (fully manual/direct control and fully autonomous operation), necessitating at least some level of human-robot interaction (HRI).

A common theme among various implementations of robotic teleoperation is the desire to move the robot to a multitude of locations within the remote environment. This must be achieved in a manner which is both efficient and safe for the operator, the robot, and the environment. The user interface (UI), through which a mobile robot is remotely controlled and monitored, can significantly affect the speed and efficacy with which these exploration and navigation tasks can be achieved. Traditional UIs often use a joystick or keyboard as input devices, with visual feedback displayed on computer monitors [12]. Immersive interfaces, such as head-mounted displays (HMDs), provide an alternative interface for robot teleoperation. It has been shown that these types of interfaces can improve the operator's situational awareness, and thus their ability to perform spatial tasks [2, 10, 15]. However, more work has to be done to explore the implications of immersive interfaces on human factors, and to ensure that they are suitable for real-world applications.

In this paper, we explore the implementation of waypoint navigation, a type of supervisory control, for a system using a HMD-based interface. In waypoint navigation, the operator selects target locations for the robot, and the robot moves to these locations autonomously. This method of control is regarded as a relatively low level of autonomy, because the operator is still responsible for all high-level mission and task planning; only local path-planning and movement is handled autonomously by the robot. For waypoint

control, the UI must provide a means by which the operator can select target positions and/or orientations for the robot. Target-selection methods used in traditional UIs generally involve clicking a target location on a map (either 2D or 3D) [11, 16, 19, 29], or on an image (or stereoscopic pair of images) from the robot's onboard camera [3, 6, 21, 35], using a computer mouse or joystick as an input device. Modern, commercial HMD systems (e.g. HTC Vive and Oculus Rift), often provide 6 degrees-of-freedom (position and orientation) tracking for the headset and gaming controllers. The tracking of the controllers allows the user to easily and accurately point to features in the virtual environment (VE) - in a similar way to how one would use a laser pointer in the real world. This function is frequently used as an input method in VR gaming, owing to its intuitiveness and effectiveness. We have leveraged this functionality to design a pointing-based target selection method for robot teleoperation, inspired by similar methods used in VR games.

This paper first explores the literature related to waypoint navigation and HMD-based interfaces in mobile robotics. We then propose a target-selection method for HMDs, based on a method used in the VR gaming industry. We then describe a VR-based user study that we designed to test the usability and effectiveness of this target selection strategy, against a baseline condition of direct control. The results of the study are reported and the implications and limitations of these results are then discussed. We conclude with a summary of our findings and recommendations for future work.

## 2 BACKGROUND

In this literature review, we focus on the user interfaces (UIs) through which waypoint navigation of mobile robots has been implemented. For a more general introduction to supervisory control methods we refer the reader to Sheridan [30].

One of the main benefits of supervisory control is improved performance under communication delays between the operator and the robot. On top of this, supervisory control can also help reduce operator workload, because the operator is no longer responsible for the lowest level of control of the robot. While the robot autonomously handles a given task (under human supervision), the operator is able to perform other functions, such as surveying the remote environment using the robot's cameras, improving their situational awareness [9]. According to Endsley [9], increased situational awareness can improve task performance and reduce the chance of errors. However, supervisory control schemes can also negatively impact HRI, especially if the operator has a low level of trust in the system, or becomes bored in their supervisory role [8]. These effects have to be considered and managed to maximise the effectiveness of the system.

Waypoint navigation, as a form of supervisory control for mobile robots, has been around for over 30 years. Some of the early work in this area came from NASA's Mars rover programme [38]. One of the challenges of designing a mobile robot for operation on Mars is the amount of time that it takes for a signal to travel between Earth and Mars. Wilcox [38] states that 'it is impractical to have a rover that is teleoperated from Earth (that is, one in which the lowest-level feedback control is mediated through real-time perception of a human being)'. The paper mentions two proposed methods

of implementing waypoint navigation for the rover; Computer-Aided Remote Driving (CARD) and semi-autonomous navigation (SAN). In CARD, the operator is presented with a pair of stereo images (using a stereoscopic display) from the robot's onboard cameras, and selects a safe path for the robot using a 3D-cursor. In SAN, global routes for the rover are plotted by the operator using a topographic map of the Mars surface. Local routes are then adapted autonomously by the rover, based on more detailed information from the onboard sensors.

In order to enable the operator to select targets in 3D space for the robot to navigate to, two primary functions have to be achieved by the interface; (1) the robot's environment has to be displayed to the operator, and (2) the operator has to be able to select a target point from that display using an input device. In terms of display, the environment is generally rendered to the operator in one of two ways; either using camera images transmitted from the robot (e.g. CARD) or using a 2D or 3D map (e.g. SAN).

Using an immersive display, such as an HMD, to render the robot's environment, rather than a computer monitor, has been shown to improve the operator's situational awareness and task performance in navigation tasks [2, 10, 26]. Some of these benefits may stem from the ability to transform some explicit controls into implicit ones; for example controlling the orientation of the camera through movement of the user's head, rather than a joystick or other control [10]. Images from the robot can also be rendered across a much larger portion of the operator's ocular field-of-view. Thus, objects in the remote environment that may appear quite small on a computer monitor, can appear much larger in a HMD. This means the operator simply has to shift their gaze to a region of interest in the rendered image, rather than manually zooming in and out of the image.

Despite these benefits, HMDs also pose additional challenges to interface design. In particular, the effects of 'simulator sickness' - namely nausea, headaches and/or vomiting - have been shown to be more severe when using HMDs, compared to a computer monitor display [28]. Moss and Muth [22] makes several recommendations to reduce these effects; (1) HMD update rate should be as high as possible (2) do not occlude the user's peripheral visual of their local environment (if possible) (3) HMD users should be provided with postural support (e.g. a railing to grasp), rather than standing freely. Another consideration is that restricting, or completely removing the operators ability to see their local environment may inhibit their ability to cooperate with other operators, and may prevent the use of certain input devices. Finally, HMDs are not nearly as familiar to the majority of users as computer monitors, and their use can induce operator stress. These issues have so far limited the use of HMDs in real world applications. More research into these interfaces is necessary in order to capitalise on the aforementioned benefits, whilst minimising these negative impacts.

Using HMD-based interfaces for waypoint navigation provides an interesting challenge. The way in which a target position is selected will be affected by the way in which the environment is being rendered to the operator. The camera images from the robot's onboard camera(s) could be streamed directly to the HMD [2, 24]. This technique restricts the viewpoint of the operator to the viewpoint of the robot's camera, i.e. an egocentric viewpoint. In this case, waypoint selection methods would have to be augmented

onto the camera images. Alternatively, the operator could be shown a 3D representation of the robot's environment. This model may be produced from *a priori* knowledge [26, 39], or 'virtualized' in real-time using 3D mapping [17, 29, 31, 32]. Rendering the environment in 3D allows the operator to separate their viewpoint from the robot's position, allowing either egocentric, exocentric or tethered viewpoints [14]. These two techniques can also be combined by projecting the camera images into the 3D environment, in the vein of the ecological interfaces developed by Nielsen et al. [23] (e.g. [1, 25, 29, 33]), or by allowing the operator to switch between the two modes. Finally, in the situation where the operator and robot are collocated, the operator may view the environment directly, with the target selection method augmented on top of the scene. This could either be achieved using a dedicated augmented-reality HMD (e.g. Microsoft HoloLens) [36] or using video-passthrough with a VR-type HMD (e.g. HTC Vive, Oculus Rift); Class 3 and Class 4 interfaces respectively in Milgram and Kishino's taxonomy [20].

The primary efforts to achieve waypoint selection with a HMD-based interface that we have found are [1, 26, 32, 36]. The target selection methods in these four implementations, which all use an exocentric viewpoint, are outlined below:

*Drag and drop.* [32] - The user generates targets for the robot by dragging "ghost versions of the robot" to the desired locations in the virtually rendered 3D world, using the tracked VR headset (HTC Vive) controllers. To facilitate this target selection method, the operator is able to scale themselves up and down, as well as translate around the VE. The VE is pre-mapped, but updated in real-time by the teleoperated unmanned ground robot (UGV). A quantitative evaluation of the system is not provided in the paper.

*Augmented reality virtual surrogate.* [36] - The operator sets waypoints for a collocated unmanned aerial vehicle (UAV), by controlling a virtual 'surrogate' UAV superimposed on the environment using an augmented-reality HMD (Microsoft HoloLens). This control method was shown to improve navigation task completion time, as well as improve the operators ability to multitask during the trial, when compared to direct control. Although this implementation relies on collocation of the operator and the robot, the same waypoint setting method could be implemented remotely if the environment was rendered virtually (as in [32]), or by superimposing the steerable 'virtual leader' robot on to the camera feed from the robot [17, 27].

*Point and click.* [1, 26] - Roldán et al. [26] proposes a pointing-based target selection method, for the control of a UGV and UAV in an outdoor multi-robot navigation task. This target selection method is essentially the same as the one described in this paper, where the operator uses the 6DoF tracked controllers to point to a target location in the virtual world. The main difference between this system and ours is that Roldán et al. [26] use an exocentric viewpoint. A full 3D virtual copy of the real environment is modelled by hand from *a priori* knowledge and presented to the operator using the HTC Vive headset. Allspaw et al. [1] describes a similar exocentric, pointing-based target selection method for the control of a humanoid robot. The authors also propose allowing the operator to place virtual footprint objects in the environment, that the robot would try to copy exactly.

As previously mentioned, exocentric viewpoints necessarily require the environment to be mapped in 3D. Therefore, robots that operate in unknown and unmapped environments would have a requirement for real-time 3D mapping. 3D maps produced by state-of-the-art real-time mapping techniques cannot replicate the level of detail of a camera image, and often require more specialist equipment. Alternatively, the camera feed from the robot's onboard camera could be streamed directly to the operator's HMD. This would allow the operator to perceive the environment in a high level of detail, but necessitates an egocentric viewpoint. Egocentric viewpoints have been shown to be preferable for local guidance [14, 37] and increased operator presence - which may make it a preferable for certain applications, such as telepresence. To this end, we propose an interface that combines pointing-based target selection with an egocentric viewpoint, in the following section. We also conducted a user study to test the efficacy of the proposed system. To the best of our knowledge, the effectiveness of interfaces that use a pointing-based target selection method (as per [26] and [1]), with an egocentric viewpoint have not been assessed.

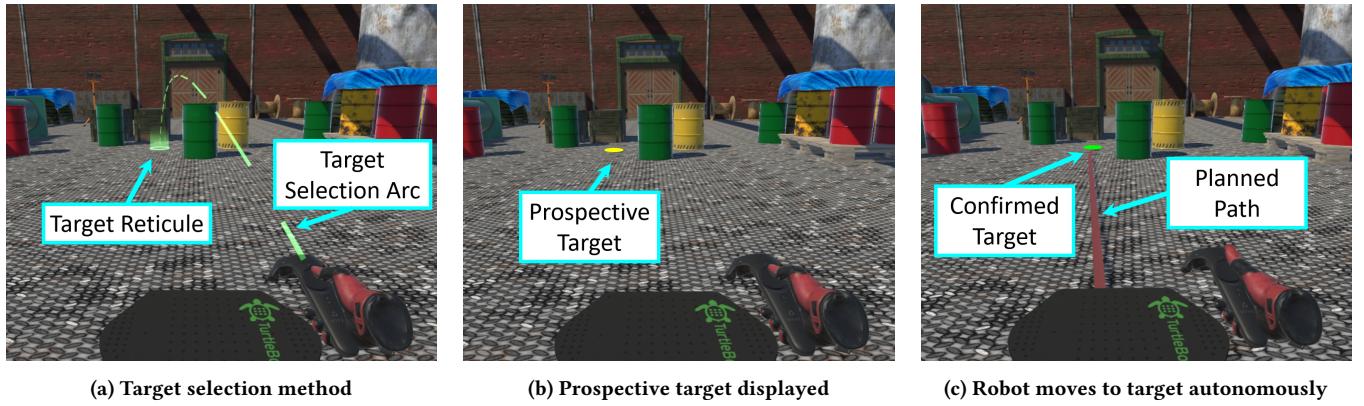
### 3 SYSTEM DESIGN

The proposed target selection strategy, is based on a method of locomotion called *teleportation* used in the VR gaming industry. Teleportation was designed to allow players to navigate around virtual environments in an intuitive manner. To perform teleportation, the player selects a point in the environment using the 6DoF tracked VR controllers as a virtual pointing device. On selection of a point, the player's position in the environment is instantly changed to the chosen position. While it is clearly not possible to teleport a robot between locations, this method does seem to be a convenient and intuitive way to select target points in VR.

The target selection we propose uses the same input method as teleportation. However, instead of transporting the user directly to the target, the position is logged and used as a goal location for the robot's onboard planner. The proposed method is shown in **Figure 1**, and in the accompanying video <sup>1</sup>. The input method works as follows: (1) The operator depresses a button on the VR controller, which causes a virtual arc to render between the controller and the ground plane (**Figure 1a**). The operator can alter the position of the target reticle by changing the position and orientation of the controllers, thereby changing the point at which the arc intersects the terrain. (2) When the operator is satisfied with the position of the target reticle, they release the button and a yellow prospective target disk is spawned at that location (**Figure 1b**). (3) The operator then confirms the selected target by pulling the trigger on the controller. The target disk changes colour to green and the robot starts planning a path to the chosen target position. Once a path has been calculated, it is communicated back to the operator, and superimposed on their display (**Figure 1c**). (4) While the robot is moving to the target, the operator may select a new target position by repeating steps (1) - (3), or can order the robot to stop moving immediately using the grip button on the controller.

One of the benefits of this target selection method is that it can work across a range of HMD-based teleoperation setups, using a standard commercial VR device (e.g. HTC Vive). Roldán et al.

<sup>1</sup>[https://youtu.be/yW6j36XRT\\_8](https://youtu.be/yW6j36XRT_8)



**Figure 1: The principle steps of the waypoint control method:** (a) target position is selected using the tracked VR controller; (b) target is displayed and confirmed by the user; (c) the robot moves autonomously to the target and publishes its path to the operator.

[26] showed that the target selection method is effective in a hand-modelled virtual environment, for a multi-robot use case. We also suggest that this method could be used for a collocated operator and robot use case (e.g. [36]), by superimposing the target-setting virtual arc onto the real environment using an ARHMD (e.g. Microsoft HoloLens). Similarly, this target-selection method is also possible when viewing the camera feed from the remote robot directly (i.e. an egocentric viewpoint), by superimposing the virtual arc onto the camera feed. This is the type of interface we are interested in developing, because it does not necessitate detailed 3D mapping of the environment.

As a first step towards this type of interface, we developed a VR testbed in which users can control a virtual robot around a virtual environment using the aforementioned target selection method. This allowed us to examine various interface design decisions in preliminary trials, ultimately culminating in a usability study on the proposed method, outlined in **Section 4**. Such experiments are important before implementation on a real robotic system, as they help ensure that the interface is viable for operators to use and provide quantitative data on which to justify interface design decisions. This should help maximise the effectiveness of the future interface and reduce the negative effects, such as minimising the VR sickness that operators feel.

## 4 METHODS

We conducted a user study to test the effectiveness of the proposed target selection method when using an egocentric viewpoint. The primary objective of the experiment was to navigate a virtual TurtleBot 2 robot from one end of an indoor environment, to a goal position at the far end. The aim of the study was to provide quantitative data and qualitative feedback on the effectiveness and usability of the interface. The experiment was run entirely in virtual reality; i.e. the robot controlled in the experiment was not a virtual ‘twin’ of a physical system operating in a real-world environment. The experiment was conducted using an HTC Vive HMD, with standard 6DoF tracked Vive controllers. The virtual environment, the robot, and the integration with the HTC Vive, were implemented using

the Unity game engine<sup>2</sup>. In this section we will first provide details of the experimental design, including the apparatus used and a description of the test environment. We also provide an outline of the experimental procedure, and identify the measures recorded during the experiment.

### 4.1 Experimental design

The waypoint navigation interface, with the proposed target selection method, was compared against a baseline direct control method. In the direct control condition, the operator used the 2-axis trackpad on the Vive controller to provide control inputs. The position of the operators thumb on the y-axis of the trackpad was used to control the forward/back motion of the robot and the position on the x-axis was converted to left/right turning motions. In this condition, a collision avoidance mechanism was implemented that prevented the robot from driving into obstacles in the environment; a function which is common in teleoperated systems. A UI warning was displayed if the proximity sensor on front of the robot was triggered and the robot was prevented from moving forward, and instead restricted to turning or reversing motions. In both setups, the operator had a full 360° view of the environment, viewable by turning their head in the VR headset. However, unlike standard VR-gaming interfaces, the spatial position of the operators viewpoint was fixed to the position of the camera on the robot, to better represent the constraints of a real-world camera system. The VR system also renders a stereoscopic pair of images to the operator each frame, allowing them to perceive depth more easily than a monocular video.

As discussed in **Section 2**, communication delays between the operator and the robot are often present in real-world systems. One of the simplifications of running a fully virtual experiment is that these communication delays are essentially negated. In order to account for this, we included an artificial communication delay in the system as a variable. In the experiment, we tested two delay conditions; no delay and a fixed 1 second delay. A fixed delay was chosen in order to minimise the possible VR sickness effects that

<sup>2</sup><https://unity.com>

the participant would experience. The 1 second duration of the delay was chosen because it is in the range at which direct control performance becomes significantly degraded [7] and aligns with the ‘short’ time delay duration used by Luck et al. [19] for a mobile robot navigation task.

A 2x2 within-subjects factorial design allowed us to test the effectiveness of the proposed interface against direct control, across two levels of delay (no delay and 1 second delay). The within-subject design allowed us to gather direct comparative feedback from the participants. The order of the control conditions was counterbalanced. However, the non-delayed condition was always run first for a given control method. This ordering was chosen so that the participant did not have to experience an unfamiliar control method and adverse delay conditions simultaneously, as it was thought that this may be too stressful and exacerbate the effects of VR sickness.

## 4.2 Experimental Procedure

During the experiment, the participant experienced the following set of events:

- (1) A short video outlining the goal of the experiment <sup>3</sup>.
- (2) A VR-based tutorial for the first control method they were going to use (direct control or waypoint navigation). During the tutorial, the participant was able to try controlling the robot with no delay and 1 second delay.
- (3) A timed navigation trial through the virtual environment, with no communication delay. The participant then answered questionnaires about the trial (see Measures).
- (4) Another timed trial, with the same control method, but with 1 second communication delay, followed by questionnaires.
- (5) A repeat of Steps (2) - (4), using the other method of control.
- (6) An optional ‘bonus’ trial, where the operator could switch between the two control methods as and when they chose (described below).
- (7) A semi-structured discussion about the experiment (see Measures).

A top down view of the virtual environment used in the timed trials is shown in **Figure 2**. The environment contained a significant number of static obstacles, including barrels, waste silos, and pillars, which supported elevated walkways. The environment was used in the forward direction (Start to Target in **Figure 2**) for the first trial (without delay), and in reverse (Target to Start) for the second trial (with delay) in order to keep the optimal path length the same. Mirrored versions of these two environments were used in the third and fourth trials (forward-mirrored and reverse-mirrored respectively) to ensure that the environment was identical between control methods for a given delay level. The potential for learning effects from reusing the same environment (albeit mirrored) was counteracted by counterbalancing the order of the control conditions. In piloting, very few participants realised that the same environment layout was being used, so learning effects due to this are likely to be minimal in any case.

In order to gauge the participant’s cognitive workload during the navigation task, the participants were also asked to perform a timed ancillary task. Left or right facing arrows would periodically

<sup>3</sup><https://youtu.be/oJFDSp7mYjw>

appear ( $\approx$  8 seconds apart) on either side of their screen, and the participant was required to click the corresponding button on the VR controller in their non-dominant hand; the dominant hand controller was being used to manoeuvre the robot. The response times to the stimuli were measured throughout the timed trials.

After the four timed trials, participants were given the option to partake in an additional ‘bonus’ round. The aim of the bonus round was to garner objective data about the participant’s preferred control method in different situations. In the bonus round, the participant was given the ability to switch between the two control methods (direct control and waypoint navigation). In the first half of the trial, no communication delay was present, but halfway through the trial a 1 second delay was imposed; the participant was notified before the trial that this was going to happen, and was alerted during the trial when it occurred. It was made known to the participants that a small monetary prize would be awarded to the participant with the lowest overall time (the sum of trial completion time and the cumulative response times to the ancillary task), in order to encourage them to optimise their task speed. In real-world mobile robot navigation missions, task speed is only one of several factors that operators may seek to optimise, but nevertheless provides a useful basis for comparison in this experiment.

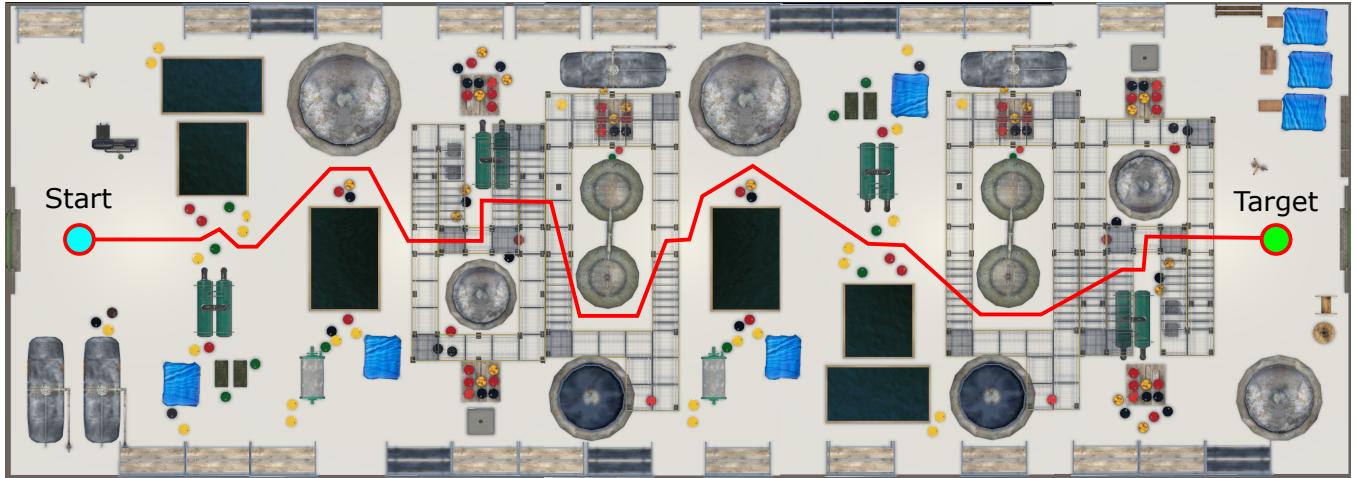
## 4.3 Measures

Our overall aim was to develop an effective and intuitive target selection method for HMD-based teleoperation. We decided to compare the effect of the proposed interface based on three factors: (a) task performance (b) cognitive workload, and (c) usability. We used both objective and subjective measures to evaluate the effects of the interface on these areas. Task performance and cognitive workload comparisons between the novel control method (waypoint control) and the familiar control method (direct control) in the non-delayed condition were chosen to show whether participants were able to pick up the novel interface quickly, thereby indicating its intuitiveness.

For task performance, primary (navigation) task completion time and total distance travelled by the robot, were used as objective measures.

For cognitive workload, response times to the ancillary (button-pressing) task and the number of ancillary task errors (i.e. pressing the wrong button) were used as objective measures. As a subjective measure, we decided to use the ‘Raw’ variant of the NASA task load index (Raw-TLX) questionnaire, where the final score is simply taken as a mean of the scores from each category: mental demand, physical demand, temporal demand, performance, effort and frustration. This method has been shown to be equally as effective as the original weighted NASA-TLX in a number of studies [13], and takes significantly less time to administer.

As a subjective measure for system usability, we utilised the system usability scale (SUS) questionnaire [5] after each trial. Qualitative feedback was obtained from the participants using a semi-structured interview at the end of the trial. During the interview participants were asked to state their preferred control method in the case of no delay, and in the case of a delay. Participants were also asked open-ended questions, such as “what were your thoughts on the two methods of control you used in this experiment?”. The



**Figure 2: Top down view of the VR trial environment, overlaid with the optimum route between the start and finish points.**

amount of time each control method was used in the bonus trial was also recorded as an objective measure of user preference.

## 5 RESULTS

### 5.1 Participants

A total of 37 participants were recruited from the university campus. 3 participants (1 male, 2 female) did not finish all 4 trials, because they experienced significant nausea effects. Data from their trials was removed, leaving 34 participants (25 male, 9 female). Demographics recorded for the participants included how often they played computer games (2 daily, 16 weekly, 12 monthly, 4 never), experience with virtual reality (1 a lot, 29 some, 4 none) and experience operating robots (12 a lot, 22 some). The average participant age was 28.2 ( $\sigma = 4.46$ ), with a range of 22-43. Out of the 34 participants that completed all 4 conditions, 22 chose to participate in the bonus round.

### 5.2 Analysis

Unless otherwise stated, the measures outlined above were analysed using a repeated-measures two-way analysis of variance (ANOVA) with control method and delay as the two independent variables, each with two levels. The order that participants experienced the two control conditions was treated as a between-subject factor, to analyse the effect that order had on the results. The error bars in the following graphs show the 95% confidence interval.

### 5.3 Objective Measures

The repeated-measures two-way ANOVA showed a significant interaction effect between control method and delay on task completion time,  $F(1, 32) = 28.86, p < 0.001$ . Order did not have a significant effect on this measure. Simple effects analysis showed that there was no significant difference between the control methods when delay was not present ( $p = 0.2$ ). However, in the delayed condition waypoint control significantly outperformed direct control,  $F(1, 32) = 57.39, p < 0.001$ . Analysis showed that there were no significant main effects of control or delay on mean response

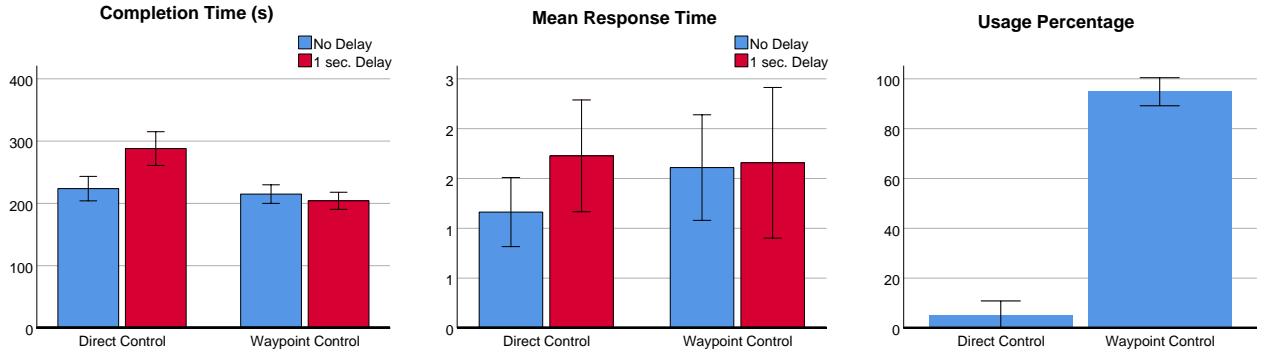
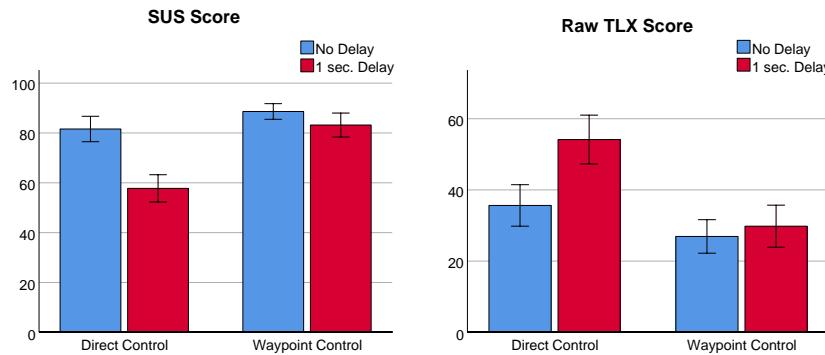
time to the secondary task. The usage percentages of each control condition during the bonus trial, where participants could switch between control methods at any time during the trial, were analysed using a paired samples t-test. The results showed a significant usage preference for waypoint control ( $M = 94.8\%$ ) over direct control ( $M = 5.2\%$ ),  $t(22) < 0.001$ .

### 5.4 Subjective Measures

Control interface usability was evaluated with the SUS questionnaire. A two-way repeated-measures ANOVA showed a significant interaction between control method and delay on the total SUS score,  $F(1, 32) = 25.889, p < 0.001$ . The analysis also showed a significant interaction between control method and order,  $F(1, 32) = 7.795, p = 0.009$ . **Table 1** shows the estimated marginal means of the SUS scores when split by order; direct control first (DC First) and waypoint control first (WC First). For the group that experienced the direct control condition first, there was no significant difference between the SUS scores in the non-delayed conditions ( $F(1, 18) = 0.25, p = 0.877$ ). However, for the group that experienced waypoint control first, the SUS scores for waypoint control were found to be significantly higher,  $F(1, 14) = 7.545, p = 0.016$ . For both groups, the SUS scores were found to be significantly higher for waypoint control in the delayed case ( $F(1, 18) = 54.724, p < 0.001$  and  $F(1, 14) = 36.977, p < 0.001$ ). When the data from both groups is combined into one set, simple effects analysis showed a significantly higher total SUS score for waypoint control in both the non-delayed condition,  $F(1, 32) = 7.21, p = 0.011$ , and the delayed condition,  $F(1, 32) = 88.89, p < 0.001$ .

The Raw-TLX was used to assess cognitive workload during the tasks. A two-way repeated-measures ANOVA showed a significant interaction between control method and delay on the total Raw-TLX score,  $F(1, 32) = 19.182, p < 0.001$ . A simple effects analysis showed a significantly lower Raw-TLX score for waypoint control in both the non delayed condition,  $F(1, 32) = 8.713, p = 0.006$ , and the delayed condition,  $F(1, 32) = 66.472, p < 0.001$ .

**Table 2** shows the significance values for the difference between waypoint and direct control on the components of the Raw-TLX.

**Figure 3: Results of objective measures. Error bars at 95% confidence intervals.****Figure 4: Results of subjective measures. Error bars at 95% confidence intervals.****Table 1: Estimated marginal means ( $M$ ) and standard deviations ( $\sigma$ ) of SUS scores when the data is split into 2 groups by order of control conditions; direct control (DC) first and waypoint control (WC) first.**

| Control | Delay  | DC First |          | WC First |          |
|---------|--------|----------|----------|----------|----------|
|         |        | $M$      | $\sigma$ | $M$      | $\sigma$ |
| DC      | None   | 86.81    | 2.01     | 77.00    | 5.05     |
|         | 1 sec. | 62.37    | 2.67     | 53.17    | 5.02     |
| WC      | None   | 86.58    | 2.34     | 90.67    | 1.80     |
|         | 1 sec. | 83.55    | 3.10     | 82.83    | 3.59     |

**Table 2: Statistical significance test results for Raw-TLX component scores. In all significant cases, waypoint control condition had lower mean values than direct control.**

| Component   | No Delay                    | 1 Second Delay               |
|-------------|-----------------------------|------------------------------|
| Mental      | Not Significant             | $F(1,32) = 29.01, p < 0.001$ |
| Physical    | $F(1,32) = 7.62, p = 0.009$ | $F(1,32) = 14.42, p = 0.001$ |
| Temporal    | $F(1,32) = 5.90, p = 0.021$ | $F(1,32) = 28.11, p < 0.001$ |
| Performance | Not Significant             | $F(1,32) = 15.91, p < 0.001$ |
| Effort      | $F(1,32) = 6.53, p = 0.016$ | $F(1,32) = 56.91, p < 0.001$ |
| Frustration | $F(1,32) = 7.14, p = 0.12$  | $F(1,32) = 50.33, p < 0.001$ |

## 6 DISCUSSION

The aim of this work is to develop an effective and intuitive target selection method for HMD-based teleoperation. The results show that the waypoint control interface used in this experiment was either similar or better than direct control in all three areas of interest (task performance, cognitive workload and usability) across both levels of delay (no delay and 1 second delay). This result comes in spite of the fact that direct control is a more familiar method of control to most users; highlighting that participants were able to pick up the target selection method quickly and thus its intuitive nature.

These results were obtained using a two-way repeated-measures ANOVAs followed by post hoc simple effects analyses.

Finally, participants were asked to state a preferred control method in a delayed and non-delayed condition. In the non-delayed condition, 2 participants (6%) stated a preference for direct control, 4 participants (12%) stated no preference, and the remaining 28 participants (82%) stated a preference for waypoint control. In the delayed condition, all participants stated a preference for waypoint control.

One overwhelming result, observed across all recorded metrics, was that the waypoint control method was significantly more effective than direct control when time delay was present. Many participants fed back that the waypoint control method was barely affected by the delay, and some (albeit few) actually preferred it with delay, stating that they felt it gave them more time to think. This result shows that our proposed interface successfully leverages one of the primary benefits of supervisory control schemes; robustness in the face of communication delay.

The results of the control method rankings and the usage percentages in the bonus round show that participants preferred using the waypoint control system, even when there was no delay, for this task. Comments from participants suggest that transferring the responsibility for local path-planning and collision avoidance to the robot allowed them to concentrate on higher level route planning. This was also reflected in the higher usability scores of the waypoint control interface, particularly in the delayed condition. Order effects were shown to be significant on usability scores; participants scored the usability of direct control significantly lower when they experienced it after waypoint control; i.e. participants felt that usability was impacted more when autonomous functionality was removed, than when it was added, which is an interesting result.

The Raw-TLX scores indicated significantly lower cognitive load when using waypoint control. However, response time to the ancillary task (the objective measure for cognitive load) did not show the same trend. In fact, in the conditions with no delay, mean response times to the secondary task appears to be lower in direct control than waypoint control (albeit not significantly). Based on observations in the trials and participant comments, the immediacy and simplicity of the secondary task meant that it was not severely affected by their cognitive workload levels. Furthermore, experimental observations indicated that waypoint control induced more fluctuation in the operators cognitive load due to the primary task. A high level of cognitive load was required while the user was selecting a target position, followed by a period of much lower cognitive load while the robot was moving autonomously. Because the secondary task required an immediate response, the results could have been affected by these short periods of high workload. Additionally, reduced focus during the autonomous manoeuvres may have also negatively affected response time. A more complex ancillary task may have provided better granularity on these cognitive workload fluctuations, and may have highlighted the differences between the two control methods more clearly. An entirely different ancillary task (e.g. being required to search for and count particular objects in the environment) might have better highlighted the effect of the participants' cognitive workload on their situational awareness, which was not explicitly measured in this experiment. Subjective workload results displayed the same ordering effects as usability, we expect for the same reasons.

VR sickness turned out to be a significant factor in this experiment, with 3 participants unable to finish all 4 conditions, and a significant proportion of participants reporting mild to moderate nausea in the post experiment interviews. Unfortunately, nausea effects were not systematically recorded as a measure in this experiment, but participant comments highlighted that there could be an impact of control method on levels of nausea experienced. The direct control condition with delay was most frequently cited

as the condition that induced the most nausea. This could be in part because of the higher level of workload and concentration required during this condition. Overall, VR sickness effects are likely to have a large impact on whether egocentric HMD-based displays are viable for mobile robot teleoperation and their effects should be measured in future experiments.

## 6.1 Limitations and Further Work

We made the decision to run the experiment virtually in order to get an indication of how an optimal implementation of this type of interface might perform. However, several additional challenges need to be tackled when applying this to a real-world system. The robot must be able to map its environment in great enough resolution for the operator to accurately select a target using the proposed method. However, because the operator uses camera images for scene understanding when using an egocentric viewpoint, the level of detail required of the 3D model should still be significantly less than that required for an exocentric implementation, where the 3D model is the operators only way of viewing the environment.

The actual communication delay present in a real-world implementation could also feasibly be shorter or longer than 1 second, and/or variable (rather than fixed), depending on factors associated with a specific setup (range of teleoperation, communication methods, camera type etc.).

When implemented on a real robot, we also intend to compare this target selection method to others, such as using a 3D joystick to place targets (e.g. [6]). We also intend to compare the effectiveness of egocentric and exocentric viewpoints. Finally, in future experiments we aim to measure nausea effects more systematically (e.g. using the Simulator Sickness Questionnaire [18]), providing insight into which types of interface cause the least VR sickness effects.

## 7 CONCLUSION

This paper has explored the utility of an egocentric HMD-based target selection interface for waypoint navigation of mobile robots. The proposed target selection method was based on the VR gaming notion of *teleportation*. We implemented this method using an HTC vive and accompanying tracked controllers to select waypoints in a virtual reality environment.

Both the effectiveness and the intuitiveness of the proposed interface were examined through a user study, involving 37 participants. The results of the VR-based user study showed that this type of target selection method is viable for implementation on real robots. Participants were able to learn how to use the system quickly, and were therefore able to leverage the beneficial characteristics of semi-autonomous control methods; namely reduced cognitive workload and mitigated communication delay effects. The experiment also highlighted that an egocentric viewpoint may be problematic for HMD-based teleoperation, because the motion of the robot's camera through the environment, viewed through the VR headset, caused significant nausea effects in some participants.

Further work will involve implementing this target selection method on a real robot, and comparing it to other methods of target selection currently used for mobile robots.

## REFERENCES

- [1] Jordan Allspaw, Lilia Heinold, and Holly A Yanco. 2019. Design of Virtual Reality for Humanoid Robots with Inspiration from Video Games. In *International Conference on Human-Computer Interaction*. 3–18.
- [2] Luis Almeida, Paulo Menezes, and Jorge Dias. 2017. Improving robot teleoperation experience via immersive interfaces. *Proceedings of 2017 4th Experiment at International Conference: Online Experimentation, exp.at 2017* (2017), 87–92. <https://doi.org/10.1109/EXPAT.2017.7984414>
- [3] Federica Bazzano, Fabrizio Lamberti, Andrea Sanna, Gianluca Paravati, and Marco Gaspardone. 2017. Comparing Usability of User Interfaces for Robotic Telepresence. In *VISIGRAPP (2: HUCAPP)*. 46–54.
- [4] Jenay M Beer, Arthur D Fish, and Wendy A Rogers. 2014. Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of human-robot interaction* 3, 2 (7 2014), 74–99. <https://doi.org/10.5898/JHRI.3.2.Beer>
- [5] John Brooke and others. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [6] Jonathan M. Cameron, Brian K. Cooper, Robert A. Salo, and Brian H. Wilcox. 1987. Fusing Global Navigation With Computer-Aided Remote Driving Of Robotic Vehicles. In *Mobile Robots I*, Vol. 0727. SPIE, 85. <https://doi.org/10.1117/12.937786>
- [7] Jessie Y C Chen, Ellen C Haas, and Michael J Barnes. 2007. Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 6 (2007), 1231–1245.
- [8] Mary L Cummings, C Mastracchio, Kristopher M Thornburg, and A Mkrtchyan. 2013. Boredom and distraction in multiple unmanned vehicle supervisory control. *Interacting with Computers* 25, 1 (2013), 34–47.
- [9] Mica R Endsley. 1988. Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society Annual Meeting* 32, 2 (1988), 97–101. <https://doi.org/10.1177/154193128803200221>
- [10] Juan C. García, Bruno Patrón, Luis Almeida, Javier Pérez, Paulo Menezes, Jorge Dias, and Pedro J. Sanz. 2017. A Natural Interface for Remote Operation of Underwater Robots. *IEEE Computer Graphics and Applications* 37, 1 (2017), 34–43. <https://doi.org/10.1109/MCG.2015.118>
- [11] Michael A Goodrich, Timothy W McLain, Jeffrey D Anderson, Jisang Sun, and Jacob W Crandall. 2007. Managing autonomy in robot teams: observations from four experiments. In *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 25–32.
- [12] D W Hainsworth. 2001. *Teleoperation User Interfaces for Mining Robotics*. Technical Report. 19–28 pages.
- [13] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. 904–908.
- [14] Justin G Hollands and Matthew Lamb. 2011. Viewpoint tethering for remotely operated vehicles: effects on complex terrain navigation and spatial awareness. *Human factors* 53, 2 (2011), 154–167.
- [15] Jarosław Jankowski and Andrzej Grabowski. 2015. Usability evaluation of vr interface for mobile robot teleoperation. *International Journal of Human-Computer Interaction* 31, 12 (2015), 882–889.
- [16] Jennifer S Kay. 1995. STRIPE: Remote driving using limited image data. In *Conference Companion on Human Factors in Computing Systems*. 59–60.
- [17] Alonso Kelly, Erin Capstick, Daniel Huber, Herman Herman, Pete Rander, and Randy Warner. 2011. Real-time photorealistic virtualized reality interface for remote mobile robot control. *Springer Tracts in Advanced Robotics* 70, STAR (2011), 211–226. [https://doi.org/10.1007/978-3-642-19457-3\\_13](https://doi.org/10.1007/978-3-642-19457-3_13)
- [18] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. [https://doi.org/10.1207/s15327108jap0303\\_3](https://doi.org/10.1207/s15327108jap0303_3)
- [19] Jason P. Luck, Patricia L. McDermott, Laurel Allender, and Deborah C. Russell. 2006. An investigation of real world control of robotic assets under communication latency. *HRI 2006: Proceedings of the 2006 ACM Conference on Human-Robot Interaction 2006*, January 2006 (2006), 202–209. <https://doi.org/10.1145/1121241.1121277>
- [20] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.
- [21] Giovanni Mosiello, Andrey Kiselev, and Amy Loutfi. 2013. Using Augmented Reality to Improve Usability of the User Interface for Driving a Telepresence Robot. *Paladyn, Journal of Behavioral Robotics* 4, 3 (2013), 174–181. <https://doi.org/10.2478/pjbr-2013-0018>
- [22] Jason D Moss and Eric R Muth. 2011. Characteristics of head-mounted displays and their effects on simulator sickness. *Human factors* 53, 3 (2011), 308–319.
- [23] Curtis W. Nielsen, Michael A. Goodrich, Robert W. Ricks, Senior Member, and Robert W. Ricks. 2007. Ecological interfaces for improving mobile robot teleoperation. *IEEE Transactions on Robotics* 23, 5 (2007), 927–941. <https://doi.org/10.1109/TRO.2007.907479>
- [24] Yeonju Oh, Ramviyas Parasuraman, Tim Mcgraw, and Byung-cheol Min. 2018. 360 VR Based Robot Teleoperation Interface for Virtual Tour. *International Workshop on Virtual, Augmented and Mixed Reality for Human-Robot Interaction* March (2018).
- [25] Tobias Rodehutsks, Max Schwarz, and Sven Behnke. 2015. Intuitive bimanual telemانipulation under communication restrictions by immersive 3D visualization and motion tracking. *IEEE-RAS International Conference on Humanoid Robots* 2015-Decem (2015), 276–283. <https://doi.org/10.1109/HUMANOIDS.2015.7363547>
- [26] JJ Roldán, E Peña-Tapia, P García-Aunon, J Del Cerro, and A Barrientos. 2019. Bringing Adaptive and Immersive Interfaces to Real-World Multi-Robot Scenarios: Application to Surveillance and Intervention in Infrastructures. *IEEE Access* 7 (2019), 86319–86335. <https://doi.org/10.1109/ACCESS.2019.2924938>
- [27] Markus Sauer, Florian Zeiger, and Klaus Schilling. 2010. Mixed-Reality User Interface for Mobile Robot Teleoperation in Ad-Hoc Networks. *IFAC Proceedings Volumes* 43, 23 (1 2010), 77–82. <https://doi.org/10.3182/20101005-4-RO-2018.00029>
- [28] Ludger Schmidt, Jens Hegenberg, and Liubov Cramar. 2014. User studies on teleoperation of robots for plant inspection. *Industrial Robot* 41, 1 (2014), 6–14. <https://doi.org/10.1108/IR-02-2013-325>
- [29] Max Schwarz, Marius Beul, David Droeßel, Sebastian Schüller, Arul Selvam Periyasamy, Christian Lenz, Michael Schreiber, and Sven Behnke. 2016. Supervised Autonomy for Exploration and Mobile Manipulation in Rough Terrain with a Centaur-Like Robot. *Frontiers in Robotics and AI* 3, October (2016). <https://doi.org/10.3389/frobt.2016.00057>
- [30] Thomas B Sheridan. 1992. *Telerobotics, automation, and human supervisory control*. MIT press.
- [31] Patrick Stotko, Stefan Krumpen, Max Schwarz, Christian Lenz, Sven Behnke, Reinhard Klein, and Michael Weinmann. 2019. A VR System for Immersive Teleoperation and Live Exploration with a Mobile Robot. *arXiv preprint arXiv:1908.02949* (2019).
- [32] Antti Tikanmaki, Tomas Bedrník, Rajesh Raveendran, and Juha Roning. 2017. The remote operation and environment reconstruction of outdoor mobile robots using virtual reality. *2017 IEEE International Conference on Mechatronics and Automation, ICMA 2017* (2017), 1526–1531. <https://doi.org/10.1109/ICMA.2017.8016043>
- [33] Karthik Mahesh Varadarajan and Markus Vincze. 2011. Augmented virtuality based immersive telepresence for control of mining robots. In *2011 5th International Symposium on Computational Intelligence and Intelligent Informatics (ISCIII)*. 133–138.
- [34] Michael A. Vidulich and Pamela S. Tsang. 2012. Mental Workload and Situation Awareness. In *Handbook of Human Factors and Ergonomics*. John Wiley & Sons, Inc., Hoboken, NJ, USA, 243–273. <https://doi.org/10.1002/9781118131350.ch8>
- [35] Richard Volpe, J Balaram, Timothy Ohm, and Robert Ivlev. 1996. The rocky 7 mars rover prototype. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'96*, Vol. 3. 1558–1564.
- [36] Michael E Walker, Hoooman Hedayati, and Daniel Szafir. 2019. Robot Teleoperation with Augmented Reality Virtual Surrogates \*. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2019), 202–210.
- [37] Christopher D Wickens, Justin G Hollands, Simon Banbury, and Raja Parasuraman. 2013. *Engineering psychology and human performance*. New York: Psychology Press. <https://doi.org/10.4324/9781315665177>
- [38] Brian H Wilcox. 1992. Robotic vehicles for planetary exploration. *Applied Intelligence* 2, 2 (1992), 181–193. <https://doi.org/10.1007/BF00058762>
- [39] Brandon Wilson, Matthew Bounds, David McFadden, Jace Regenbrecht, Loveth Ohenen, Alireza Tavakkoli, and Donald Loffredo. 2018. VETO: An Immersive Virtual Environment for Tele-Operation. *Robotics* 7, 2 (2018), 26.